

Compute-Service am RZ

Erste Gedanken zum Compute-Service wurden im Rahmen des SERVUZ-Projektes entwickelt. Sie haben mit dem Kauf einer Convex Metaserie, bestehend aus einer Convex C3820 ES und sechs HP-PA-RISC-Knoten (Workstations HP 9000/735), Gestalt angenommen.

Die Convex-Serie C38xx besteht aus Hochleistungsrechnern mit integrierter Vektor- und Parallelverarbeitung, die durch die Gallium-Arsenid-Technik hohe Zuverlässigkeit und Geschwindigkeit in einer luftgekühlten Umgebung bieten. Die C3820 ES ist ein Einstiegsmodell in diese höchste Leistungsklasse von Convex-Supercomputern. Sie besitzt zwei Prozessoren, die über einen Crossbarswitch mit dem virtuell verwalteten Hauptspeicher verbunden sind. Das Peripheriesubsystem besteht aus 2 VME I/O-Prozessoren (VIOP), siehe auch Tabelle 1.

Die PA-RISC-Knoten sind keine Workstations im herkömmlichen Sinne, da sie über kein Display und kein Keyboard verfügen. Sie sind in einem Convex-Kabinett untergebracht. Die Konsolports der Knoten sind mit RS-232-Ports der Convex direkt verbunden, um eine Administration von der Convex-Konsole aus

zu ermöglichen. Der Datenaustausch zwischen den PA-RISC-Knoten und der C3820 erfolgt über ein separates Netz - gegenwärtig über Ethernet, ab Juli über einen in das Kabinett der PA-RISC-Knoten integrierten FDDI-Konzentrator.

Die genannte Hardware, Convex C3820 ES und HP-Cluster, bilden mit den genannten Verbindungen und den weiter unten beschriebenen Softwarekomponenten einen Ansatz für einen sogenannten Metacomputer. Dieser Begriff wurde von Larry Smarr und Charles Catlett geprägt, die darunter ein Netzwerk von heterogenen Computern, die durch Hard- und Software verbunden sind, verstehen. Dem Nutzer erscheint das Netzwerk als eine einzige Compute-Ressource. Das Ziel der Metaserie besteht darin, Daten- und CPU-intensive Aufgaben zu trennen und auf die geeigneten Rechner zu verteilen sowie eine parallele Abarbeitung zu ermöglichen. Die C3820 bietet die Vorteile einer hohen I/O-Bandbreite, eines großen Hauptspeichers sowie einer Vektorverarbeitung gegenüber der hohen skalaren Geschwindigkeit der PA-RISC-Prozessoren (siehe Tabelle 1).

	CONVEX C3820 ES	HP 9000/735
Prozessoren	2 CPUs, 2 VIOPs	PA-RISC 7100
Hauptspeicher	512 MB	32 MB
Plattenspeicher	16 x 2,5 GB Diskarray	1 GB
Cache		256 MB
Taktfrequenz		99 MHz
Prozessorzykluszeit	16,67 ns	
Peak	480 MFLOPS	40 MFLOPS
FDDI	DAS (dual attached station)	SAS (single attached station)
externe Datenträger	2 GB DAT-Laufwerk	CD ROM-Laufwerk

Tabelle 1: Hardwareparameter des Compute-Service

Clusterkonzept:

In einer Metaserie können entweder alle integrierten Rechner zu einem physischen Cluster zusammengefaßt werden, oder es können logische Cluster gebildet werden, die nur Teilmengen der heterogenen Rechner umfassen. Die C3820 fungiert als Systemmastermaschine, die mittels geeigneter Software alle Systemfunktionen auch auf den PA-RISC-Knoten ausführen kann. Die Cluster werden durch den Systemadministrator vordefiniert. Alle Kommandos beziehen sich auf die vordefinierten Cluster oder können durch eine Option -C in der Kommandozeile oder durch die Umgebungsvariable CLUSTER auf den benannten Cluster bezogen werden.

Der Compute-Service liefert eine Umgebung für maximalen Durchsatz und eine für Parallelverarbeitung. Ein Cluster kann für maximalen Durchsatz oder/und für Parallelverarbeitung konfiguriert werden.

ConvexNQS+:

Im Falle, daß ein Cluster für maximalen Durchsatz konfiguriert wurde, wird eine hohe skalare Compute-Leistung geliefert, um Batch-Jobs zu bearbeiten. Als Software dient das ConvexNQS+, das auf der C3820 sowie auf allen PA-RISC-Knoten installiert ist. ConvexNQS+ unterstützt zwei Typen von Queues. Die erste ist eine reine Batch-Queue. Die zweite ist eine Pipe-Queue, die Jobs weiterleitet an eine weitere Pipe-Queue oder eine Batch-Queue desselben oder

eines anderen Hosts. Eine Pipe-Queue routet, basierend auf einer Zielliste, entweder zur ersten Queue, die fähig ist, die Anforderung zu empfangen, oder zur ersten Queue, die fähig ist, den Request sofort auszuführen oder zur Queue mit der leichtesten Last (Last ist eine Kombination aus durchschnittlicher Systemlast, Prozessorgeschwindigkeit, Queuelänge und einem Wichtungsfaktor). Das ConvexNQS+ unterstützt Ressourcenlimits auf der C3820 und auf den PA-RISC-Knoten, wobei auf letzteren CPU-Limits, Core_file_size-Limits, Data_file_size-Limits, Data_segment_size-Limits und Working_set-Limits nicht unterstützt werden. ConvexNQS+ unterstützt Checkpointing und Restart von Batch-Requests.

ConvexPVM:

Im Falle der Konfiguration für Parallelverarbeitung liefert die Metaserie einen grobkörnigen Parallelismus, der durch Messagepassing zwischen den Programmmoduln erreicht wird. Mittels ConvexPVM (Convex Version von Parallel Virtual Machine) erfolgt das Messagepassing und das Abschieken von Jobs, die auf mehreren Hosts gleichzeitig rechnen sollen. Das ConvexPVM behandelt heterogene Rechnersysteme wie Komponenten der gleichen virtuellen Maschine. Der Nutzer kann Systemrufe des ConvexPVM sowohl in C- als auch in FORTRAN-Programme einbetten. Diese signalisieren der Convex das Entladen von hochparallelisierten Algorithmen auf die PA-RISC-Knoten und gestatten somit das vorrangige Ausführen solcher Routinen auf der C3820, die deren Vektorarchitektur und I/O-Bandbreite erfordern. In ConvexPVM sind mit PVMdb auch debugging Eigenschaften integriert. ConvexMLIB ist eine auf CONVEX VECLIB basierende Bibliothek von Routinen, die mathematische Software und Berechnungen für Anwendungsprogramme, die Felder einschließen, liefert. Sie ist optimiert für PA-RISC und läuft auf der C3820 und den PA-RISC-Knoten.

Filesysteme:

Die Nutzerfilesysteme des gesamten Compute-Service liegen generell auf der Convex. Sie können mittels NFS an den PA-RISC-Knoten gemounted werden. Das Nutzerfilesystem /home wird mittels Quotas überwacht. Die PA-RISC-Knoten werden nur Filesysteme für Betriebssystem, Swapping, Zusatzsoftware und temporäre Dateien haben. Die Datensicherung für die Metaserie wird über die File-Service-Funktionen der C3820 (s. nachfolgenden Beitrag) erfolgen.

Systemsoftware:

CONVEX C3820 ES	HP 9000/735
ConvexOS (~ BSD 4.3)	HP/UX 9.0
Fortran Compiler	Fortran Compiler
C Compiler	C Compiler
NFS	NFS
TCP/IP	TCP/IP
CXwindows	
OSF/Motif	
Consultant III (Debugger)	
Fiar Share Scheduler	
Remote System Backup	
VECLIB	
ConvexNQS+	ConvexNQS+
ConvexPVM	ConvexPVM
ConvexMLIB	ConvexMLIB
UniTree	

Nutzersoftware:

(1) Nutzersoftware, die bisher auf anderen Rechnern zur Verfügung stand, wird auch auf der C3820 ES installiert.

Bibliotheken:

- NAG Mark 14 (Fortran-Bibliothek)
- CERN-Bibliotheken

Chemie:

- AMBER (molecular Modelling)
- DISCOVER (" " , zu Insight auf SGI)
- GAUSSIAN 92 (Quanten-Chemie)

(2) Nutzersoftware, die noch in diesem Jahr gekauft werden soll.

Compiler:

- Pascal-Compiler (Universität Tübingen)

Mathematik/Statistik:

- SAS (Statistikpaket)
- Mathematica (symbol. Rechnen)

Visualisierung:

- AVS (Application Visualization System)

Chemie:

- SCHAKAL (Universität Freiburg)
- VAMP 4.4 (Quanten-Chemie)
- SPARTAN 2.1 (ab initio Rechnungen, Graphik auf SGI)

Daniela-Maria Subklew